

Asymptotic distribution of unbiased linear estimators in the presence of heavy-tailed stochastic regressors and residuals*

JEONG-RYEOL KURZ-KIM, SVETLOZAR T. RACHEV AND
GENNADY SAMORODNITSKY

Abstract

Under the symmetric α -stable distributional assumption for the disturbances, Blattberg et al (1971) consider unbiased linear estimators for a regression model with non-stochastic regressors. We consider both the rate of convergence to the true value and the asymptotic distribution of the normalized error of the linear unbiased estimators. By doing this, we allow the regressors to be stochastic and disturbances to be heavy-tailed with either finite or infinite variances, where the tail-thickness parameters of the regressors and disturbances may be different.

Keywords: Asymptotic distribution, rate of convergence, stochastic regressor, stable non-Gaussian, finite or infinite variance, heavy tails

*The views expressed in this paper are those of the authors and not necessarily those of the Deutsche Bundesbank or other institutions involved. Kurz-Kim gratefully acknowledge research support from the Alexander von Humboldt Foundation. Rachev's research was supported by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara and the Deutschen Forschungsgemeinschaft. Samorodnitsky's research was partially supported by the Deutschen Forschungsgemeinschaft and by NSA grant MSPF-02G-183 and NSF grant DMS-0071073 at Cornell University.

1 Introduction

For the estimation of the coefficients of a regression model one typically applies ordinary least squares (OLS), which is equivalent to the maximum likelihood estimation if the disturbances are normally distributed. Furthermore, according to the Gauss-Markov theorem, the OLS estimator has the minimum variance of all linear unbiased estimators if the disturbances follow a distribution with finite variance. However, if the disturbances follow a distribution with infinite variance, but with finite mean, the OLS estimator is still unbiased but no longer a minimum variance estimator.

Relaxing the normality assumption by allowing disturbances to have a symmetric α -stable distribution with infinite variance ($1 < \alpha < 2$), Blattberg et al (1971) generalize the OLS estimator to a different linear unbiased estimator that minimizes the α -stable scale of the estimator. That generalization is performed in the framework of a regression model in which the independent variable is assumed to be non-stochastic.

We consider both the rate of convergence to the true value and the asymptotic distribution of the normalized error of the linear unbiased estimators of coefficients in the regression model with both stochastic regressors and disturbances being heavy-tailed with either finite or infinite variances,¹ and the tail-thickness parameters of the regressors and disturbances may be different. Even though our distributional assumptions are more general than the assumptions of α -stability, the *limiting distributions* of the estimators will often be expressed through stable random variables.

For any random variable X there is a number $\alpha \in (0, 2]$ satisfying $C^\alpha = A^\alpha + B^\alpha$. Exponent α is called the *stability parameter*. A random variable with exponent α is said to be α -stable distributed. Closed-form expressions of α -stable distributions exist only for a few special cases. However, the logarithm of the characteristic function of the α -stable distribution can be written as (see Zolotarev (1986) and Samorodnitsky et al (1994) for more details on α -stable distributions)

$$\ln \varphi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha [1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}] + i\mu t, & \text{for } \alpha \neq 1, \\ -\sigma |t| [1 + i\beta \frac{\pi}{2} \operatorname{sign}(t) \ln |t|] + i\mu t, & \text{for } \alpha = 1, \end{cases}$$

where α is the *stability parameter* (or tail-thickness parameter); σ is the *scale pa-*

¹There is some controversy on whether the variance of financial returns is always infinite. We avoid this controversy by using a heavy-tailed model that allows for both finite or infinite variance.

parameter; $\beta \in [-1, 1]$ is the *skewness parameter*; and μ is the *location parameter*. If $\beta = 0$, the distribution is symmetric. The shape of the symmetric α -stable distribution ($S\alpha S$) is determined by the tail-thickness parameter $\alpha \in (0, 2]$. For $0 < \alpha < 2$ the tails of the distribution are thicker than those of the normal distribution; and the tail-thickness increases as α decreases. When $\alpha = 2$, the $S\alpha S$ distribution coincides with the normal distribution with variance $2\sigma^2$, the only member of the family with finite variance. When $\alpha = 1$, the $S\alpha S$ distribution reduces to the Cauchy distribution. If $\alpha < 2$, moments of order α or higher do not exist, which means the variance is infinite. If X is an α -stable random variable, $0 < \alpha < 2$, with scale σ , skewness β , and location μ , then a common notation is $X \sim S_\alpha(\sigma, \beta, \mu)$. In that case the tails of X are given by

$$P(\pm X > \lambda) \sim C_\alpha \frac{1 \pm b}{2} \sigma^\alpha \lambda^{-\alpha} \quad (1)$$

as $\lambda \rightarrow \infty$, where

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x \, dx \right)^{-1}. \quad (2)$$

Some more basic information and notation on stable random variables we use, unless otherwise specified, can be found in Samorodnitsky et al (1994).

One distinct example for a possible application of our results in the paper can be found in financial market analysis. For econometric analysis of a dynamic capital asset pricing model, it is necessary to add an assumption concerning the distributional behavior of stock returns. Since Bachelier (1900) the traditional and most widely adopted distributional assumption on financial return process has been the Gaussian assumption. Due to the influential works of Mandelbrot (1963) and Fama (1965), however, the α -stable distributions with $0 < \alpha < 2$ have often been considered to be a more realistic distribution assumption for asset returns than that of a normal distribution, because asset returns are typically fat-tailed and excessively peaked around zero—phenomena that can be captured by α -stable distributions with $\alpha < 2$. This is the so-called *stable Paretian assumption*. In a certain sense the stable Paretian assumption is a generalization rather than an alternative to the Gaussian assumption. Indeed, according to the generalized central limit theorem, the limiting distribution of the sum of a large number of independent, identically distributed (*iid*) random variables is α -stable with $0 < \alpha \leq 2$; see Zolotarev (1986). For more applications of

the α -stable distributions in economics and finance; see Rachev et al (1999), Rachev et al (2000) and Kim (2000).

The rest of the paper is organized as follows. In Section 2 we present our new estimator and analyze the asymptotic distribution of the estimator. Section 3 summarizes various scenarios by the different α for regressors and disturbances. In Section 4 we discuss the choice of the optimal θ in the new estimator both analytically and numerically. In Section 5, using simulation and response surface analysis, we present both the limiting and finite-sample distributions of our new estimator. Section 6 contains some concluding remarks.

2 Rate of convergence and the limiting distribution for the regression coefficient estimator

Consider a simple regression model like the one below.

$$Y_j = \beta X_j + U_j, \quad j = 1, 2, \dots \quad (3)$$

We assume that the regressors $\{X_j\}$ are *iid* random variables with polynomially decaying tails. Specifically,

$$P(|X_1| > \lambda) \sim D_1 \lambda^{-\alpha_x}, \quad \lambda \rightarrow \infty, \quad \text{some } \alpha_x > 0 \text{ and } D_1 > 0. \quad (4)$$

Furthermore, we assume that the noise (disturbances) $\{U_j\}$ are also *iid* random variables, which we assume to be symmetric, with

$$P(|U_1| > \lambda) \sim D_2 \lambda^{-\alpha_u}, \quad \lambda \rightarrow \infty, \quad \text{some } \alpha_u > 0 \text{ and } D_2 > 0. \quad (5)$$

We assume, further, that the sequences $\{X_j\}$ and $\{U_j\}$ are independent.

Note that no assumptions on the symmetry of either dependent observations or regressors are made. We remark, further, that it is relatively straightforward (at least, away from the boundary cases) to extend the results below to the case where the tails of the regressors and noise variables are regularly varying (ie adding slowly varying factors in (4) and (5)). Since such slowly varying functions are not practically observable, we decided against including extra technical arguments in an already

highly technical paper. Finally, we allow values of α_x and α_u in the interval $(0, 1]$ as well, since our methods cover those cases equally well.

The goal is to estimate the regression coefficient β in (3), and our estimator is

$$\hat{\beta}_{\theta,n} = \frac{\sum_{j=1}^n X_j^{<1/(\theta-1)>} Y_j}{\sum_{j=1}^n |X_j|^{\theta/(\theta-1)}} \quad (6)$$

for some $\theta > 1$ with $< \cdot >$ defined as a signed power.² Note that the OLS estimator corresponds to $\theta = 2$ in (6).

Our immediate task is to understand the behavior of the difference

$$\begin{aligned} \Delta_n := \hat{\beta}_{\theta,n} - \beta &= \frac{\sum_{j=1}^n X_j^{<1/(\theta-1)>} U_j}{\sum_{j=1}^n |X_j|^{\theta/(\theta-1)}} \\ &\stackrel{d}{=} \frac{\sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{\sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}, \end{aligned} \quad (7)$$

where the last distributional equality follows from the symmetry of the noise. That is, we are interested in the rate of convergence of the estimator $\hat{\beta}_{\theta,n}$ to the true value depending on the choice of θ . When such convergence actually takes place, this will also establish consistency (in probability) of our estimator.

It is clear that the rate of convergence to zero of the difference Δ_n depends significantly on the tail exponents α_x and α_u , and on the choice of θ . What is interesting is that we will see below that there are basically 7 different cases of possible values of α_x and α_u , in each of which the rate of convergence is a different function of θ .

A common feature of our results will be the existence of an exponent d such that

$$n^d \Delta_n \Rightarrow W \quad \text{as } n \rightarrow \infty \quad (8)$$

for some non-degenerate weak limit W . Occasionally, on certain boundaries we will have to modify (8) to allow for a slowly varying factor in the left-hand side. That is, we will have

$$n^d L(n) \Delta_n \Rightarrow W \quad \text{as } n \rightarrow \infty \quad (9)$$

where L is a slowly varying function.³ In any case we will view the exponent d in either (8) or (9) as measuring the rate of convergence. In particular, the exponent d

² $a^{<p>} = |a|^{p-1}a$.

³ $L(x)$ is a slowly varying function as $x \rightarrow \infty$, if for every constant $c > 0$ and $\lim_{x \rightarrow \infty} \frac{L(cx)}{L(x)}$ exists and is equal to 1.

turns out to be a different function of θ in the 7 different cases of possible values of α_x and α_u we mentioned above.

The reader will find it easier to follow the different technical detail below after noticing the existence of several critical boundaries. The first boundary is that

$$\frac{\theta - 1}{\theta} \alpha_x > 1. \quad (10)$$

Note that on one side of that boundary $|X_j|^{\theta/(\theta-1)}$ has a finite mean and hence the denominator in (7) is governed by the law of large numbers (LLN). On the other side of that boundary $|X_j|^{\theta/(\theta-1)}$ is in the domain of attraction of a positive stable law and the corresponding heavy-tailed central limit theorem governs the behavior of the denominator in (7). On the boundary itself, the mean is infinite, but the (weak) LLN is still in force.

The second critical boundary is that of

$$\min((\theta - 1)\alpha_x, \alpha_u) > 2. \quad (11)$$

Here on one side of the boundary the random variables $|X_j|^{1/(\theta-1)}U_j$ have a finite variance and hence the Gaussian central limit theorem (CLT) governs the behavior of the numerator in the second expression in (7). On the other side of that boundary these random variables are in the domain of attraction of a symmetric stable non-Gaussian law and hence the corresponding CLT will be responsible for the behavior of the numerator. On the boundary itself the variance is infinite, but the CLT will still be in force.

We now proceed to consider the different ranges of α_x and α_u mentioned above.

Scenario 1 Suppose that

$$0 < \alpha_x \leq 1 \quad \text{and} \quad \alpha_u \geq 2. \quad (12)$$

Note that under this scenario (10) fails independently of θ . Consider, therefore, the second part of the critical boundary (11)

$$(\theta - 1)\alpha_x \geq 2 \quad \text{or, equivalently,} \quad \theta \geq \frac{2 + \alpha_x}{\alpha_x}. \quad (13)$$

We will see that, in this case, the exponent d governing the rate of convergence of Δ_n to zero in (8) or (9) is given by

$$d = \frac{2\theta - (\theta - 1)\alpha_x}{2(\theta - 1)\alpha_x}. \quad (14)$$

Consider first the non-boundary case

$$\alpha_u > 2 \quad \text{and} \quad \theta > \frac{2 + \alpha_x}{\alpha_x}. \quad (15)$$

Then, we claim that the following version of (8)

$$n^d \Delta_n \Rightarrow \frac{(EU_1^2)^{1/2} (E|X_1|^{2/(\theta-1)})^{1/2}}{C_{\alpha_x(\theta-1)/\theta} D_1^{\theta/\alpha_x(\theta-1)}} \frac{N(0,1)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1,1,0)} \quad (16)$$

holds weakly, where $N(0,1)$ and $S_{\frac{(\theta-1)\alpha_x}{\theta}}(1,1,0)$ in the right-hand side above are independent. Furthermore, D_1 is the constant in the tail in (4). See the end of the previous section for the description of the constant C_α for $0 < \alpha < 2$ (given in (2)) as well as for other basic information on stable random variables.

Indeed, in this case,

$$n^d \Delta_n = \frac{n^{-1/2} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-\theta/(\theta-1)\alpha_x} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}. \quad (17)$$

Let $\varepsilon > 0$, and

$$K_n(\varepsilon) = \{j = 1, 2, \dots, n : |X_j| > \varepsilon n^{1/\alpha_x}\}. \quad (18)$$

Note that

$$k_n(\varepsilon) := \text{Card}(K_n(\varepsilon)) \Rightarrow \mathbb{P}\text{ois}(D_1 \varepsilon^{-\alpha_x}) \quad \text{as } n \rightarrow \infty \quad (19)$$

weakly, where $\mathbb{P}\text{ois}(\mu)$ stands for a Poisson random variable with mean μ .

Write (by giving names to the numerator and denominator in the right hand side of (17))

$$n^d \Delta_n = \frac{NU_n}{DE_n}, \quad (20)$$

and let

$$\hat{N}_n = n^{-1/2} \sum_{j=1}^n |\hat{X}_j|^{1/(\theta-1)} U_j, \quad (21)$$

$$\hat{D}_n = n^{-\theta/(\theta-1)\alpha_x} \sum_{j \in K_n(\varepsilon)} |X_j|^{\theta/(\theta-1)}, \quad (22)$$

where

$$\hat{X}_j = \begin{cases} X_j, & \text{if } j = \{1, 2, \dots, n\} \setminus K_n(\varepsilon) \\ \tilde{X}_j, & \text{if } j \in K_n(\varepsilon). \end{cases} \quad (23)$$

In (23) $\{\tilde{X}_j\}$ is an *iid* sequence with a common law $\mathbb{P}(X_j \in \cdot \mid |X_j| \leq \varepsilon n^{1/\alpha_x})$, and independent of the sequences $\{X_j\}$ and $\{U_j\}$.

Note that \hat{N}_n and \hat{D}_n are independent. By the CLT for triangular arrays we see that

$$\hat{N}_n \Rightarrow N(0, E[|X_1|^{2/(\theta-1)}] E[U_1^2])) \quad (24)$$

(eg Theorem 5.1.2 in Laha et al (1979)).

Furthermore, we claim that

$$\hat{D}_n \Rightarrow \sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon), \quad (25)$$

where N_ε is $\text{Pois}(D_1 \varepsilon^{-\alpha_x})$, independent of an *iid* sequence $\{Z_j(\varepsilon)\}$ with a common law

$$\mathbb{P}(Z_j(\varepsilon) > \lambda) = \left(\frac{\lambda^{(\theta-1)/\theta}}{\varepsilon} \right)^{-\alpha_x}, \quad \lambda \geq \varepsilon^{\theta/(\theta-1)}. \quad (26)$$

This is, however, clear because of (19) and the fact that

$$\mathbb{P}\left(n^{-\theta/(\theta-1)\alpha_x} |X_1|^{\theta/(\theta-1)} > \lambda \mid |X_1| > \varepsilon n^{1/\alpha_x}\right) \Rightarrow \mathbb{P}(Z_1(\varepsilon) > \lambda), \quad n \rightarrow \infty \quad \forall \lambda. \quad (27)$$

We conclude that

$$\frac{\hat{N}_n}{\hat{D}_n} \Rightarrow \left(E[|X_1|^{2/(\theta-1)}] E[U_1^2] \right)^{1/2} \frac{N(0, 1)}{\sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon)}, \quad (28)$$

with the numerator and the denominator on the right-hand side of (28) being independent.

Note that

$$\begin{aligned} E \left[e^{-\theta \sum_{j=1}^{N_\varepsilon} Z_j(\varepsilon)} \right] &\rightarrow \exp \left\{ -D_1 \int_0^\infty (1 - e^{-\theta x}) \frac{(\theta-1)\alpha_x}{\theta} x^{-(\theta-1)\alpha_x/\theta-1} dx \right\} \\ &= E \left[\exp \left\{ -\theta C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)} S_{(\theta-1)\alpha_x/\theta}(1, 1, 0) \right\} \right] \end{aligned} \quad (29)$$

for $\theta > 0$ as $\varepsilon \rightarrow 0$. Therefore, (16) will follow once we show that for all $\delta > 0$

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{NU_n}{DE_n} - \frac{\hat{N}_n}{\hat{D}_n} \right| > \delta \right) = 0; \quad (30)$$

see Theorem 3.2 in Billingsley (1999). To this end, it is enough to prove that

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{NU_n}{DE_n} - \frac{\hat{N}_n}{DE_n} \right| > \delta \right) = 0, \quad (31)$$

and

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{\hat{N}_n}{\hat{D}_n} - \frac{\hat{N}_n}{DE_n} \right| > \delta \right) = 0. \quad (32)$$

We will start with (31). Since $(\frac{1}{DE_n})$ is tight, it is enough to prove that for every $\delta > 0$

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbf{P} \left(|N_n - \hat{N}_n| > \delta \right) = 0. \quad (33)$$

We have

$$\begin{aligned} |NU_n - \hat{N}_n| &= n^{-1/2} \left| \sum_{j \in K_n(\varepsilon)} \left(|X_j|^{1/(\theta-1)} - |\tilde{X}_j|^{1/(\theta-1)} \right) U_j \right| \\ &\leq n^{-1/2} \sum_{j \in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j + n^{-1/2} \sum_{j \in K_n(\varepsilon)} |\tilde{X}_j|^{1/(\theta-1)} U_j, \end{aligned} \quad (34)$$

and so (31) will follow once we show that for all $\delta > 0$

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbf{P} \left(n^{-1/2} \sum_{j \in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j > \delta \right) = 0, \quad (35)$$

$$\lim_{n \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbf{P} \left(n^{-1/2} \sum_{j \in K_n(\varepsilon)} |\tilde{X}_j|^{1/(\theta-1)} U_j > \delta \right) = 0. \quad (36)$$

Note that as $n \rightarrow \infty$

$$\begin{aligned} \mathbf{E} \left(n^{-1/2} \sum_{j \in K_n(\varepsilon)} |X_j|^{1/(\theta-1)} U_j \right)^2 &= n (\mathbf{E}[k_n(\varepsilon)]) \mathbf{E}[U_1^2] \mathbf{E}[|X_1|^{2/(\theta-1)} |X_1| > \varepsilon n^{1/\alpha_x}] \\ &= \frac{n}{\mathbf{P}(|X_1| > \varepsilon n^{1/\alpha_x})} \mathbf{E}[k_n(\varepsilon)] \mathbf{E}[U_1^2] \mathbf{E}[|X_1|^{2/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x})] \\ &\sim C_1^{-1} \varepsilon^{-\alpha_x} D_1 \varepsilon^{-\alpha_x} \mathbf{E}[U_1^2] \cdot 0 \\ &= 0, \end{aligned} \quad (37)$$

and (35) follows. The proof of (36) is similar and even easier. Hence we have established (31).

We now switch to proving (32). Since (\hat{N}_n) , $(\frac{1}{\hat{D}_n})$ and $(\frac{1}{DE_n})$ are all tight, it is enough to prove that

$$\lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \mathbf{P} \left(|DE_n - \hat{D}_n| > \delta \right) = 0. \quad (38)$$

Notice that

$$DE_n - \hat{D}_n = n^{-\theta/((\theta-1)\alpha_x)} \sum_{j=\{1,2,\dots,n\} \setminus K_n(\varepsilon)} |X_j|^{\theta/(\theta-1)} \quad (39)$$

and hence

$$DE_n - \hat{D}_n \Rightarrow_{n \rightarrow \infty} \int_0^{\varepsilon^{\theta/(\theta-1)}} x N_*(dx), \quad (40)$$

where

$$N_* = \sum_{j=1}^{\infty} \delta_{\{D_1^{\theta/((\theta-1)\alpha_x)} \Gamma_j^{-\theta/((\theta-1)\alpha_x)}\}}, \quad (41)$$

is the appropriate Poisson random measure. Here (Γ_j) represents the arrival times of a unit rate homogeneous Poisson process on $(0, \infty)$. See, e.g., problem 4.4.2.8 in Resnick (1987). Since

$$\int_0^{\varepsilon^{\theta/(\theta-1)}} x N_*(dx) \rightarrow_{\varepsilon \rightarrow 0} 0 \quad \text{a.s.}, \quad (42)$$

we have established (38), and so have proved (32). That completes the proof of (16).

Consider now the boundary case

$$\alpha_u = 2 \quad \text{and} \quad \theta > \frac{2 + \alpha_x}{\alpha_x}. \quad (43)$$

With d still given by (14), this time we have the following version of (9):

$$\frac{n^d}{(\log n)^{1/2}} \Delta_n \Rightarrow \frac{\left(D_2 \mathbb{E}[|X_1|^{2/(\theta-1)}]\right)^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)} D_1^{\theta/\alpha_x(\theta-1)} S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)} \frac{N(0, 1)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)}, \quad (44)$$

the random variables on the right-hand side of (44) being, once again, independent. The proof is similar to that of (16) above, but instead of the CLT for triangular arrays with a finite variance it uses the general CLT for triangular arrays as in, for example, Theorem 5.3.2 in Laha et al (1979). Recall that here D_2 is the tail constant in (5).

The second boundary case

$$\alpha_u > 2 \quad \text{and} \quad \theta = \frac{2 + \alpha_x}{\alpha_x} \quad (45)$$

is similar. In that case d still given by (14) and

$$\frac{n^d}{(\log n)^{1/2}} \Delta_n \Rightarrow \frac{(D_1 \mathbb{E}[|U_1|^2])^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)} D_1^{\theta/\alpha_x(\theta-1)} S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)} \frac{N(0, 1)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)}, \quad (46)$$

where the random variables on the right-hand side of (46) are independent. Finally, in the boundary case

$$\alpha_u = 2 \quad \text{and} \quad \theta = \frac{2 + \alpha_x}{\alpha_x} \quad (47)$$

we find out, in a similar manner, that d is given by (14) and

$$\frac{n^d}{\log n} \Delta_n \Rightarrow \frac{(D_1 D_2)^{1/2}}{C_{\alpha_x(\theta-1)/\theta}^{-\theta/\alpha_x(\theta-1)} D_1^{\theta/\alpha_x(\theta-1)} S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)} \frac{N(0, 1)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)}, \quad (48)$$

and the random variables on the right-hand side of (48) are independent. The reason for the extra power of the logarithm is that in this boundary case

$$\mathbb{P}\left(|X_1|^{2/(\theta-1)}U_1^2 > y\right) \sim D_1 D_2 \frac{\log y}{y}$$

as $y \rightarrow \infty$.

Consider now the other side of the critical boundary (11)

$$(\theta - 1)\alpha_x < 2 \quad \text{or, equivalently,} \quad \theta < \frac{2 + \alpha_x}{\alpha_x}. \quad (49)$$

Here, the exponent d turns out to be given by

$$d = \frac{1}{\alpha_x}, \quad (50)$$

and we will see that the following version of (8) holds:

$$n^d \Delta_n \Rightarrow D_1^{-1/\alpha_x} \frac{\sum_{j=1}^{\infty} \Gamma_j^{-1/((\theta-1)\alpha_x)} U_j}{\sum_{j=1}^{\infty} \Gamma_j^{-\theta/((\theta-1)\alpha_x)}} \quad (51)$$

Here, as usual, (Γ_j) represents the arrival times of a unit rate homogeneous Poisson process on $(0, \infty)$, independent of the sequence (U_j) .

Note that, unlike the limits in (16) and its versions (44), (46) and (48), the numerator and denominator on the right-hand side of (51) are NOT independent (and, of course, the numerator in the latter expression is no longer a Gaussian random variable but, rather, a symmetric stable random variable with parameter of stability equal to $(\theta - 1)\alpha_x$.)

Indeed, here

$$n^d \Delta_n = \frac{n^{-1/((\theta-1)\alpha_x)} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-\theta/((\theta-1)\alpha_x)} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}.$$

Let $\varepsilon > 0$ and write

$$\begin{aligned} n^d \Delta_n &= \frac{\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(|n^{-1/\alpha_x} X_j| > \varepsilon) U_j}{\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)}} \\ &\quad + \frac{\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(|n^{-1/\alpha_x} X_j| \leq \varepsilon) U_j}{\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)}} \\ &:= M_n(\varepsilon) + R_n(\varepsilon), \quad n = 1, 2, \dots \end{aligned} \quad (52)$$

Note that

$$M_n(\varepsilon) \stackrel{d}{=} \frac{\sum_{i=1}^{N_n((\varepsilon, \infty))} K_{i,n}^{1/(\theta-1)} U_i}{\int_0^\infty x^{\theta/(\theta-1)} N_n(dx)}, \quad n = 1, 2, \dots, \quad (53)$$

where

$$N_n = \sum_{j=1}^n \delta_{\{n^{-1/\alpha_x} |X_j|\}}, \quad n = 1, 2, \dots, \quad (54)$$

and $K_{1,n} \geq K_{2,n} \dots \geq K_{n,n}$ are the size-ordered points of N_n .

Recalling (see, once again, Resnick (1987)) that

$$N_n \Rightarrow \sum_{j=1}^{\infty} \delta_{\{D_1^{1/\alpha_x} \Gamma_j^{-1/\alpha_x}\}} := N, \quad \text{as } n \rightarrow \infty \quad (55)$$

weakly in $[-\infty, +\infty] \setminus \{0\}$, we see that

$$\begin{aligned} M_n(\varepsilon) &\Rightarrow \frac{\sum_{i=1}^{N((\varepsilon, \infty))} K_i^{1/(\theta-1)} U_i}{\int_0^{\infty} x^{\theta/(\theta-1)} N(dx)} \\ &= D_1^{-1/\alpha_x} \frac{\sum_{j=1}^{\infty} \Gamma_j^{-1/((\theta-1)\alpha_x)} \mathbf{1}(\Gamma_j^{-1/\alpha_x} > \varepsilon) U_j}{\sum_{j=1}^{\infty} \Gamma_j^{-\theta/((\theta-1)\alpha_x)}} \\ &:= L(\varepsilon), \end{aligned} \quad (56)$$

weakly as $n \rightarrow \infty$, where (K_i) stands for the size-ordered points of N . Note that, almost surely,

$$\begin{aligned} L(\varepsilon) &\xrightarrow{\varepsilon \rightarrow 0} D_1^{-1/\alpha_x} \frac{\sum_{j=1}^{\infty} \Gamma_j^{-1/((\theta-1)\alpha_x)} U_j}{\sum_{j=1}^{\infty} \Gamma_j^{-\theta/((\theta-1)\alpha_x)}} \\ &:= L, \end{aligned} \quad (57)$$

the right-hand side of (51). Therefore, an appeal to Theorem 3.2 in Billingsley (1999) shows that, to prove the latter, it remains to be demonstrated that for any $\lambda > 0$

$$\lim_{\varepsilon \rightarrow \infty} \overline{\lim}_{n \rightarrow \infty} \mathbb{P}(|R_n(\varepsilon)| > \lambda) = 0. \quad (58)$$

Clearly the sequence $\{(\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)})^{-1}\}$ is (asymptotically) tight. Given $\delta > 0$ we can choose $M > 0$ and n_0 such that

$$\mathbb{P} \left(\sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{\theta/(\theta-1)} \leq M \right) \leq \delta, \quad \text{all } n \geq n_0. \quad (59)$$

Then for all $n \geq n_0$ and $\lambda > 0$

$$\begin{aligned} &\mathbb{P}(|R_n(\varepsilon)| > \lambda) \leq \delta \\ &+ \mathbb{P} \left(\left| \sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x} |X_j| \leq \varepsilon) U_j \right| > \lambda M \right). \end{aligned} \quad (60)$$

For $K > 0$ we have

$$\begin{aligned}
& \mathbb{P} \left(\left| \sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1}(n^{-1/\alpha_x} |X_j| \leq \varepsilon) U_j \right| > \lambda M \right) \\
& \leq \mathbb{P} \left(\left| \sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1} \left(n^{-1/\alpha_x} |X_j| \leq \varepsilon, n^{-1/\alpha_x} |X_j| |U_j|^{\theta-1} \leq K \right) U_j \right| > \lambda M \right) \\
& \quad + n \mathbb{P} \left(|X_1| |U_1|^{\theta-1} > K n^{1/\alpha_x} \right).
\end{aligned} \tag{61}$$

Keeping K fixed, we have by the symmetry, using the equivalence of different moments of Bernoulli random variables (see eg Proposition 3.4.1 in Kwapień et al (1992)), also known as the Khinchine inequalities,

$$\begin{aligned}
& \mathbb{P} \left(\left| \sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1} \left(n^{-1/\alpha_x} |X_j| \leq \varepsilon, n^{-1/\alpha_x} |X_j| |U_j|^{\theta-1} \leq K \right) U_j \right| > \lambda M \right) \\
& \leq \frac{1}{\lambda M} \mathbb{E} \left[\left| \sum_{j=1}^n |n^{-1/\alpha_x} X_j|^{1/(\theta-1)} \mathbf{1} \left(n^{-1/\alpha_x} |X_j| \leq \varepsilon, n^{-1/\alpha_x} |X_j| |U_j|^{\theta-1} \leq K \right) U_j \right|^2 \right] \\
& \leq c n^{-1/((\theta-1)\alpha_x)} \mathbb{E} \left(\sum_{j=1}^n |X_j|^{2/(\theta-1)} \mathbf{1} \left(|X_j| \leq \varepsilon n^{1/\alpha_x}, |X_j| |U_j|^{\theta-1} \leq K n^{1/\alpha_x} \right) U_j^2 \right)^{1/2} \\
& \leq c n^{1/2-1/((\theta-1)\alpha_x)} \left(\mathbb{E} \left(|X_1|^{2/(\theta-1)} U_1^2 \mathbf{1} \left(|X_1| \leq \varepsilon n^{1/\alpha_x}, |X_1| |U_1|^{\theta-1} \leq K n^{1/\alpha_x} \right) \right) \right)^{1/2}.
\end{aligned}$$

Here and in the sequel c is an arbitrary finite and positive constant that does not have to be the same every time it appears. By the assumption (49) we have

$$\begin{aligned}
& \mathbb{E} \left(|X_1|^{2/(\theta-1)} U_1^2 \mathbf{1} \left(|X_1| \leq \varepsilon n^{1/\alpha_x}, |X_1| |U_1|^{\theta-1} \leq K n^{1/\alpha_x} \right) \right) \\
& \sim a(\varepsilon) (n^{1/\alpha_x})^{2/(\theta-1)-\alpha_x} \\
& := a(\varepsilon) n^\rho, \quad n \rightarrow \infty
\end{aligned}$$

where $a(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$, so that the right-hand side of (61) is

$$\leq c (a(\varepsilon))^{1/2} + n \mathbb{P} \left(|X_1| |U_1|^{\theta-1} > K n^{1/\alpha_x} \right).$$

since $\alpha_u \geq 2$. Now (58) follows after letting $K \rightarrow \infty$ (we are using, once again, the fact that $\alpha_u \geq 2$) and so we have proved (51).

Scenario 2 Suppose that

$$0 < \alpha_x \leq 1, \quad 0 < \alpha_u < 2 \quad \text{and} \quad \alpha_u \geq \alpha_x. \tag{62}$$

We are now on one side of both critical boundaries (10) and (11), and the different ranges of θ appear here depending on which of the two elements under the minimum in (11) is smaller. Consider first the range

$$(\theta - 1)\alpha_x \geq \alpha_u \quad \text{or, equivalently,} \quad \theta \geq \frac{\alpha_u}{\alpha_x} + 1. \quad (63)$$

In this case it turns out that the exponent d is given by

$$d = \frac{\alpha_u \theta - (\theta - 1)\alpha_x}{(\theta - 1)\alpha_u \alpha_x}. \quad (64)$$

Again, we start with a non-boundary case

$$\theta > \frac{\alpha_u}{\alpha_x} + 1. \quad (65)$$

We claim that here

$$n^d \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (E[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)}} \frac{S_{\alpha_u}(1, 0, 0)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)} \quad (66)$$

weakly, with the random variables on the right-hand side of (66) being independent.

Indeed, here

$$n^d \Delta_n = \frac{n^{-1/\alpha_u} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-\theta/((\theta-1)\alpha_x)} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}. \quad (67)$$

The proof is parallel to that of (16). We use the notation of (18), (19), (20) and (22), while instead of (21) we use, obviously,

$$\hat{N}_n = n^{-1/\alpha_u} \sum_{j=1}^n |\hat{X}_j|^{1/(\theta-1)} U_j, \quad (68)$$

with $\{\hat{X}_j\}$ given by (23). In particular, (25) still holds. We will show now that

$$\hat{N}_n \Rightarrow C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (E[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u} S_{\alpha_u}(1, 0, 0) \quad (69)$$

weakly as $n \rightarrow \infty$. Since by the CLT,

$$N_n \Rightarrow C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (E[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u} S_{\alpha_u}(1, 0, 0) \quad (70)$$

weakly as $n \rightarrow \infty$ (see eg Chapter XVII in Feller (1966)), (69) will follow if we check that

$$\hat{N}_n - N_n \rightarrow_{n \rightarrow \infty} 0 \quad \text{in probability.} \quad (71)$$

Now,

$$\begin{aligned}\hat{N}_n - N_n &= n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j (|\hat{X}_j|^{1/(\theta-1)} - |X_j|^{1/(\theta-1)}) \\ &= n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |\hat{X}_j|^{1/(\theta-1)} - n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)}.\end{aligned}\quad (72)$$

Hence, (71) will follow once we prove that

$$n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)} \rightarrow_{n \rightarrow \infty} 0 \quad \text{in probability.} \quad (73)$$

$$n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |\hat{X}_j|^{1/(\theta-1)} \rightarrow_{n \rightarrow \infty} 0 \quad \text{in probability.} \quad (74)$$

Consider (73). Let $0 < p < 1 \wedge \alpha_u$. This gives us

$$\begin{aligned}& \mathbb{E} \left[\left| n^{-1/\alpha_u} \sum_{j \in K_n(\varepsilon)} U_j |X_j|^{1/(\theta-1)} \right|^p \right] \\ & \leq n^{-p/\alpha_u} \cdot \mathbb{E}[k_n(\varepsilon)] \mathbb{E}[U_1]^p \mathbb{E} \left[|X_1|^{p/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x}) \right] \\ & = n^{-p/\alpha_u} \left(n \mathbb{P}(|X_1| > \varepsilon n^{1/\alpha_x}) \right) \mathbb{E}[U_1]^p \\ & \quad \left(\frac{1}{\mathbb{P}(|X_1| > \varepsilon n^{1/\alpha_x})} \mathbb{E} \left[|X_1|^{p/(\theta-1)} \mathbf{1}(|X_1| > \varepsilon n^{1/\alpha_x}) \right] \right) \\ & \sim c n^{-p/\alpha_u+1} n^{-1+p/((\theta-1)\alpha_x)} \\ & = c n^{-p(1/\alpha_u-1/((\theta-1)\alpha_x))} \rightarrow_{n \rightarrow \infty} 0.\end{aligned}$$

Hence (73) holds, and the proof of (74) is the same, but easier. The rest of the proof of (66) is the same as that of (16) above.

In the boundary case

$$\theta = \frac{\alpha_u}{\alpha_x} + 1. \quad (75)$$

the exponent d is still given by (66), and the convergence statement is

$$\frac{n^d}{(\log n)^{1/\alpha_u}} \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} (D_1 D_2)^{1/\alpha_u}}{C_{(\theta-1)\alpha_x/\theta}^{-\theta/((\theta-1)\alpha_x)} D_1^{\theta/((\theta-1)\alpha_x)}} \frac{S_{\alpha_u}(1, 0, 0)}{S_{\frac{(\theta-1)\alpha_x}{\theta}}(1, 1, 0)} \quad (76)$$

weakly, the random variables on the right-hand side above still being independent.

The proof of (76) is the same as that of (66), except that (70) is now replaced by

$$n^{-1/\alpha_u} (\log n)^{-1/\alpha_u} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j \Rightarrow C_{\alpha_u}^{-1/\alpha_u} (D_1 D_2)^{1/\alpha_u} S_{\alpha_u}(1, 0, 0) \quad (77)$$

weakly as $n \rightarrow \infty$ (see Feller (1966)).

The second possibility for the minimum in (11) is the range

$$(\theta - 1)\alpha_x < \alpha_u \quad \text{or, equivalently,} \quad \theta - 1 < \frac{\alpha_u}{\alpha_x} + 1. \quad (78)$$

Here, once again,

$$d = \frac{1}{\alpha_x},$$

and (51) still holds, with the same argument.

Scenario 3 Suppose that

$$0 < \alpha_u \leq 1 \quad \text{and} \quad \alpha_u < \alpha_x. \quad (79)$$

and consider the following 3 ranges for θ :

$$\theta \geq \frac{\alpha_x}{\alpha_x - \alpha_u}, \quad (80)$$

$$\frac{\alpha_u}{\alpha_x} + 1 \leq \theta < \frac{\alpha_x}{\alpha_x - \alpha_u}, \quad (81)$$

and

$$\theta > \frac{\alpha_u}{\alpha_x} + 1. \quad (82)$$

We claim that under (80) Δ_n does not converge in probability to 0. That is, in the case that the estimator (6) is not consistent.

Indeed, let d be given by (64), and notice that now $d \leq 0$. Since the reciprocal of the fraction on the right-hand side of (67) is clearly tight, we see that Δ_n cannot converge to zero (it is not even tight if $\theta > \frac{\alpha_x}{\alpha_x - \alpha_u}$.)

Consider now the θ range (81). Here d is the same as in (64):

$$d = \frac{\alpha_u \theta - (\theta - 1)\alpha_x}{(\theta - 1)\alpha_u \alpha_x}.$$

Now d is positive, and, in the non-boundary case, (66) still holds, with the same argument, while on the (only) boundary (76) holds in the same way.

Finally, we consider the range (82). Here, once again,

$$d = \frac{1}{\alpha_x},$$

and (51) holds.

Scenario 4 Suppose that

$$1 < \alpha_x \leq 2 \quad \text{and} \quad \alpha_u \geq 2. \quad (83)$$

The first range of θ we consider is on one side of the boundary (10)

$$\theta \geq \frac{\alpha_x}{\alpha_x - 1}. \quad (84)$$

Here the exponent d turns out to be

$$d = \frac{1}{2}. \quad (85)$$

Specifically, in the non-boundary case

$$\alpha_u > 2 \quad \text{and} \quad \theta > \frac{\alpha_x}{\alpha_x - 1}, \quad (86)$$

we actually have

$$n^d \Delta_n \Rightarrow \frac{(\mathbb{E}[|X_1|^{2/(\theta-1)}])^{1/2} (\mathbb{E}[U_1^2])^{1/2}}{\mathbb{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1) \quad (87)$$

weakly as $n \rightarrow \infty$. Indeed, here

$$n^d \Delta_n = \frac{n^{-1/2} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-1} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}, \quad (88)$$

and the strong LLN applies in the denominator, while the CLT for *iid* random variables with a finite variance applies in the numerator.

Let us look at the boundary cases. If

$$\alpha_u = 2 \quad \text{and} \quad \theta > \frac{\alpha_x}{\alpha_x - 1} \quad (89)$$

then (87) is modified to

$$\frac{n^d}{(\log n)^{1/2}} \Delta_n \Rightarrow \frac{(D_2 \mathbb{E}[|X_1|^{2/(\theta-1)}])^{1/2}}{\mathbb{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1) \quad (90)$$

by using the general CLT for *iid* random variables in the numerator (see eg Proposition 5.3.3 in Laha et al (1979)).

In the second boundary case

$$\alpha_u > 2, \alpha_x > 2 \quad \text{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \quad (91)$$

the CLT for *iid* random variables with a finite variance still applies in the numerator in (88), but the finite law means LLN no longer holds in the denominator. Instead, we will use the weak LLN

$$\frac{1}{n \log n} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)} \rightarrow D_1 \quad (92)$$

in probability as $n \rightarrow \infty$; see Theorem VII.7.2 in Feller (1966). This leads to the convergence result

$$(n^d \log n) \Delta_n \Rightarrow \frac{(E[|X_1|^{2/(\theta-1)}])^{1/2} (E[U_1^2])^{1/2}}{D_1} N(0, 1) \quad (93)$$

weakly as $n \rightarrow \infty$.

The next boundary case is

$$\alpha_u > 2, \quad \alpha_x = 2 \quad \text{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1}. \quad (94)$$

Here we use both the general CLT for *iid* random variables and the weak LLN (94) and obtain

$$n^d (\log n)^{1/2} \Delta_n \Rightarrow \frac{(E[|U_1|^2])^{1/2}}{D_1^{1/2}} N(0, 1). \quad (95)$$

Similarly, in the boundary case

$$\alpha_u = 2, \quad \alpha_x > 2 \quad \text{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \quad (96)$$

we have

$$n^d (\log n)^{1/2} \Delta_n \Rightarrow \frac{(D_2 E[|X_1|^{2/(\theta-1)}])^{1/2}}{D_1} N(0, 1). \quad (97)$$

Finally, in the boundary case

$$\alpha_u = 2, \quad \alpha_x = 2 \quad \text{and} \quad \theta = \frac{\alpha_x}{\alpha_x - 1} \quad (98)$$

we will have

$$n^d \Delta_n \Rightarrow \frac{D_2^{1/2}}{D_1^{1/2}} N(0, 1); \quad (99)$$

see the discussion after (48).

The next range of θ we consider is on the other side of the boundary (10), but still on the same side of the boundary (11)

$$\frac{2 + \alpha_x}{\alpha_x} \leq \theta < \frac{\alpha_x}{\alpha_x - 1}. \quad (100)$$

The exponent d here is given by (14):

$$d = \frac{2\theta - (\theta - 1)\alpha_x}{2(\theta - 1)\alpha_x}$$

and, in fact, in the non-boundary case (15), the convergence result (16) holds, with the same argument as before, whereas in the three boundary cases (43), (45) and (47), we obtain the convergence results (44), (46) and (48) respectively, once again with the same argument as in Scenario 1.

Finally, consider the range of θ on the other side of the boundary (11):

$$\theta < \frac{2 + \alpha_x}{\alpha_x}. \quad (101)$$

Here, the exponent d is given by (50):

$$d = \frac{1}{\alpha_x},$$

and as above, the convergence result (51) holds.

Scenario 5 Suppose that

$$\alpha_x > 1, \quad 1 < \alpha_u < 2 \quad \text{and} \quad \alpha_u \leq \frac{\alpha_x}{\alpha_x - 1}. \quad (102)$$

Once again, we start with θ on one side of the critical boundary (10):

$$\theta \geq \frac{\alpha_x}{\alpha_x - 1}. \quad (103)$$

Here it turns out that

$$d = 1 - \frac{1}{\alpha_u}. \quad (104)$$

Consider first the non-boundary case

$$\theta > \frac{\alpha_x}{\alpha_x - 1}. \quad (105)$$

Since

$$n^d \Delta_n = \frac{n^{-1/\alpha_u} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-1} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}, \quad (106)$$

we can use the CLT in the numerator and the LLN in the denominator to obtain immediately that

$$n^d \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (E[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{E[|X_1|^{\theta/(\theta-1)}]} S_{\alpha_u}(1, 0, 0) \quad (107)$$

weakly as $n \rightarrow \infty$.

In the boundary case

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad \text{and} \quad \alpha_u < \frac{\alpha_x}{\alpha_x - 1}, \quad (108)$$

we can still the CLT in the numerator, but this time we need to use the weak LLN (92) in the denominator, and obtain

$$(n^d \log n) \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} D_2^{1/\alpha_u} (E[|X_1|^{\alpha_u/(\theta-1)}])^{1/\alpha_u}}{D_1} S_{\alpha_u}(1, 0, 0) \quad (109)$$

weakly as $n \rightarrow \infty$. Furthermore, in the second boundary case

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad \text{and} \quad \alpha_u = \frac{\alpha_x}{\alpha_x - 1}, \quad (110)$$

and with the same treatment of the denominator, we need to use the version of the CLT given in (77) to obtain

$$n^d (\log n)^{1-1/\alpha_u} \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} (D_1 D_2)^{1/\alpha_u}}{D_1} S_{\alpha_u}(1, 0, 0) \quad (111)$$

weakly as $n \rightarrow \infty$.

In the second range of θ we consider, as we are now on the other side of the critical boundary (10),

$$\frac{\alpha_u}{\alpha_x} + 1 \leq \theta < \frac{\alpha_x}{\alpha_x - 1} \quad (112)$$

(note that this range is non-empty only if $\alpha_u < \alpha_x/(\alpha_x - 1)$.) Here, we take

$$d = \frac{\alpha_u \theta - (\theta - 1) \alpha_x}{(\theta - 1) \alpha_u \alpha_x}. \quad (113)$$

In fact, the convergence results obtained here are (66) in the non-boundary case (65), and (76) in the boundary case (75), all with the same arguments as above.

In the final range of θ we preserve the side of the critical boundary (10) we are on, but the relationship between the heaviness of the tails of the random variables under the sum defining the numerator of Δ_n in the second equality in (7) changes:

$$\theta < \frac{\alpha_u}{\alpha_x} + 1. \quad (114)$$

Here the exponent d is again given by (50),

$$d = \frac{1}{\alpha_x},$$

and, as before, we obtain the convergence result (51).

Scenario 6 Suppose that

$$\alpha_x > 2 \quad \text{and} \quad \alpha_u \geq 2. \quad (115)$$

The first range of θ we consider puts us on one side of the critical boundary (11):

$$\theta \geq \frac{2 + \alpha_x}{\alpha_x}. \quad (116)$$

Here d is given by (85):

$$d = \frac{1}{2}.$$

Specifically, in the non-boundary case

$$\theta > \frac{2 + \alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u > 2 \quad (117)$$

we have the convergence result (87). In the boundary case

$$\theta > \frac{2 + \alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u = 2 \quad (118)$$

we have the weak convergence in (90), and in the boundary case

$$\theta = \frac{2 + \alpha_x}{\alpha_x} \quad \text{and} \quad \alpha_u = 2 \quad (119)$$

we have

$$\frac{n^d}{\log n} \Delta_n \Rightarrow \frac{(D_1 D_2)^{1/2}}{\mathbb{E}[|X_1|^{\theta/(\theta-1)}]} N(0, 1) \quad (120)$$

weakly as $n \rightarrow \infty$. Once again, see the discussion after (48).

The second range of θ we consider puts us on the other side of the critical boundary (11), but keeps us on the same side of the critical boundary (10):

$$\frac{\alpha_x}{\alpha_{x-1}} \leq \theta < \frac{2 + \alpha_x}{\alpha_x}. \quad (121)$$

It turns out that in this case the exponent d is given by

$$d = \frac{(\theta - 1)\alpha_x - 1}{(\theta - 1)\alpha_x}. \quad (122)$$

Consider first the non-boundary case

$$\theta > \frac{\alpha_x}{\alpha_{x-1}} \quad (123)$$

and observe that here

$$n^d \Delta_n = \frac{n^{-1/(\theta-1)\alpha_x} \sum_{j=1}^n |X_j|^{1/(\theta-1)} U_j}{n^{-1} \sum_{j=1}^n |X_j|^{\theta/(\theta-1)}}. \quad (124)$$

By using the CLT for sums of *iid* random variables with a finite variance in the numerator and the LLN for *iid* random variables with a finite mean in the denominator we immediately obtain

$$\begin{aligned} n^d \Delta_n &\Rightarrow \frac{C_{(\theta-1)\alpha_x}^{-1/((\theta-1)\alpha_x)} D_1^{1/((\theta-1)\alpha_x)} (\mathbb{E}[|U_1|^{(\theta-1)\alpha_x}])^{1/((\theta-1)\alpha_x)}}{\mathbb{E}[|X_1|^{\theta/(\theta-1)}]} \\ &\times S_{(\theta-1)\alpha_x}(1, 0, 0) \end{aligned} \quad (125)$$

weakly as $n \rightarrow \infty$.

In the boundary case

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad (126)$$

we use, similarly, the general weak LLN in the denominator to obtain

$$\begin{aligned} (n^d \log n) \Delta_n &\Rightarrow \frac{C_{(\theta-1)\alpha_x}^{-1/((\theta-1)\alpha_x)} D_1^{1/((\theta-1)\alpha_x)} (\mathbb{E}[|U_1|^{(\theta-1)\alpha_x}])^{1/((\theta-1)\alpha_x)}}{D_1} \\ &\times S_{(\theta-1)\alpha_x}(1, 0, 0) \end{aligned} \quad (127)$$

weakly as $n \rightarrow \infty$.

The final range of θ puts us on the other side of the critical boundary (10):

$$\theta < \frac{\alpha_x}{\alpha_x - 1}. \quad (128)$$

Here d is still given by (50):

$$d = \frac{1}{\alpha_x}$$

and the weak convergence (51) still holds.

Scenario 7 Suppose that

$$\alpha_x > 2 \quad \text{and} \quad \frac{\alpha_x}{\alpha_x - 1} < \alpha_u < 2. \quad (129)$$

The first range of θ specifies which of the two elements under the minimum in (11) is smaller:

$$\theta \geq \frac{\alpha_u}{\alpha_x} + 1. \quad (130)$$

Here d turns out to be still given by (104):

$$d = 1 - \frac{1}{\alpha_u}.$$

In the non-boundary case

$$\theta > \frac{\alpha_u}{\alpha_x} + 1. \quad (131)$$

and the weak convergence in (107) still holds. In the boundary case

$$\theta = \frac{\alpha_u}{\alpha_x} + 1 \quad (132)$$

the usual appeal to the general CLT for *iid* summands gives us

$$\frac{n^d}{(\log n)^{1/\alpha_u}} \Delta_n \Rightarrow \frac{C_{\alpha_u}^{-1/\alpha_u} (D_1 D_2)^{1/\alpha_u}}{\mathbb{E}[|X_1|^{\theta/(\theta-1)}]} S_{\alpha_u}(1, 0, 0) \quad (133)$$

weakly as $n \rightarrow \infty$.

The next range for θ changes which of the two elements under the minimum in (11) is smaller, but still keeps us on the same side of the critical boundary (10):

$$\frac{\alpha_x}{\alpha_x - 1} \leq \theta < \frac{\alpha_u}{\alpha_x} + 1. \quad (134)$$

Here, d is given by (122):

$$d = \frac{(\theta - 1)\alpha_x - 1}{(\theta - 1)\alpha_x}.$$

In the non-boundary case

$$\theta > \frac{\alpha_x}{\alpha_x - 1} \quad (135)$$

the weak convergence in (125) still holds, and with the same argument. In the boundary case

$$\theta = \frac{\alpha_x}{\alpha_x - 1} \quad (136)$$

the weak convergence in (127) holds.

The final range for θ puts us on the other side of the critical boundary (10):

$$\theta < \frac{\alpha_x}{\alpha_x - 1}. \quad (137)$$

Here d is given by (50):

$$d = \frac{1}{\alpha_x},$$

and we have the weak convergence in (51).

3 Summary of different scenarios

In this section we summarize the seven possible scenarios considered above. We start with a plot showing how the scenarios partition the positive quadrant.

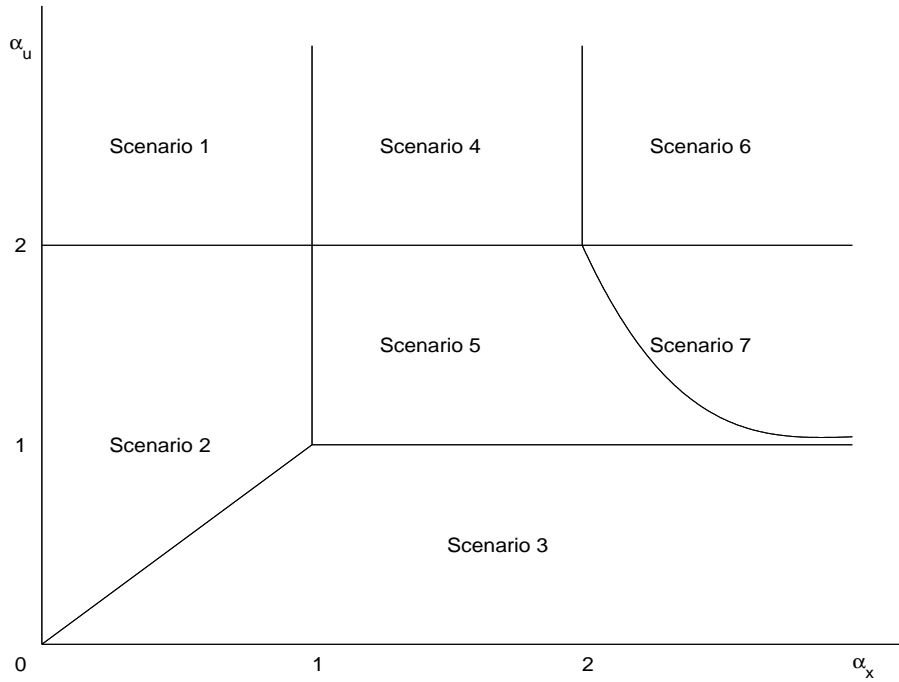


Figure 1: All possible scenarios

Recall that the exponent d describes the rate of convergence of the estimator (6); see (8) and (9). Under each one of the seven scenarios this exponent is a different function of the parameter θ .

In the sequel we look at each scenario separately and state the behavior of the exponent $d = d(\theta)$ for $\theta > 1$.

Scenario 1: $0 < \alpha_x \leq 1$ and $\alpha_u \geq 2$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{2+\alpha_x}{\alpha_x} \\ \frac{2\theta - (\theta-1)\alpha_x}{2(\theta-1)\alpha_x} & \text{if } \theta > \frac{2+\alpha_x}{\alpha_x} \end{cases}.$$

See the plot below.

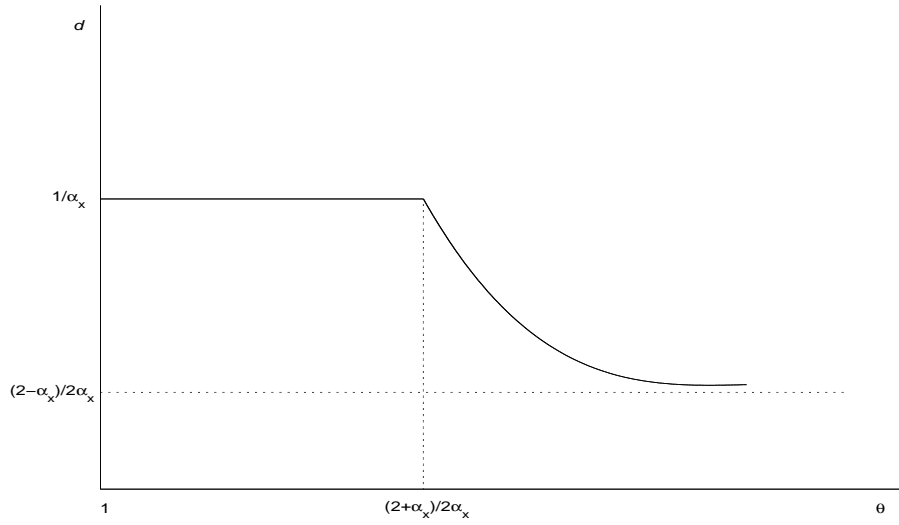


Figure 2: Scenario 1

Scenario 2: $0 < \alpha_x \leq 1$, $0 < \alpha_u < 2$ and $\alpha_u \geq \alpha_x$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{\alpha_u}{\alpha_x} + 1 \\ \frac{\theta\alpha_u - (\theta-1)\alpha_x}{(\theta-1)\alpha_u\alpha_x} & \text{if } \theta > \frac{\alpha_u}{\alpha_x} + 1 \end{cases}.$$

See the plot below.

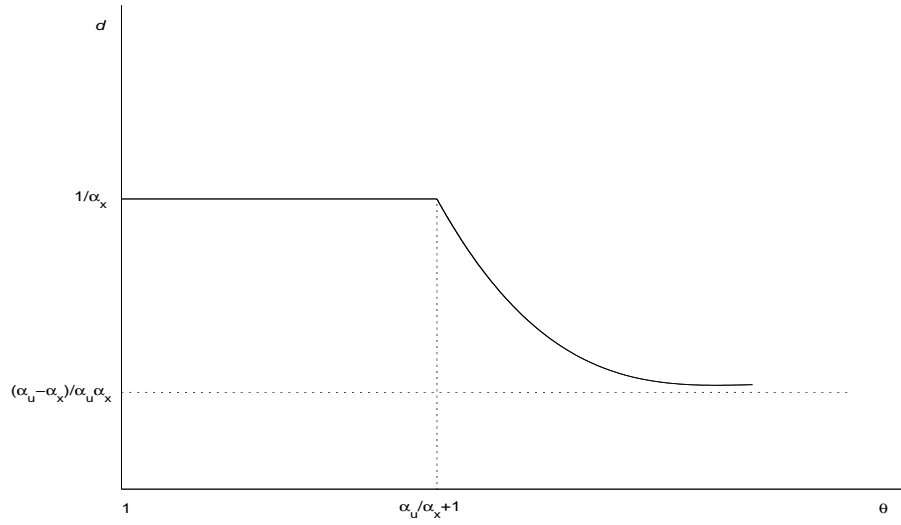


Figure 3: Scenario 2

Scenario 3: $0 < \alpha_u \leq 1$ and $\alpha_u < \alpha_x$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{\alpha_u}{\alpha_x} + 1 \\ \frac{\theta\alpha_u - (\theta-1)\alpha_x}{(\theta-1)\alpha_u\alpha_x} & \text{if } \frac{\alpha_u}{\alpha_x} + 1 < \theta < \frac{\alpha_x}{\alpha_x - \alpha_u} \\ \text{no consistency} & \text{if } \theta \geq \frac{\alpha_x}{\alpha_x - \alpha_u} \end{cases}.$$

See the plot below.

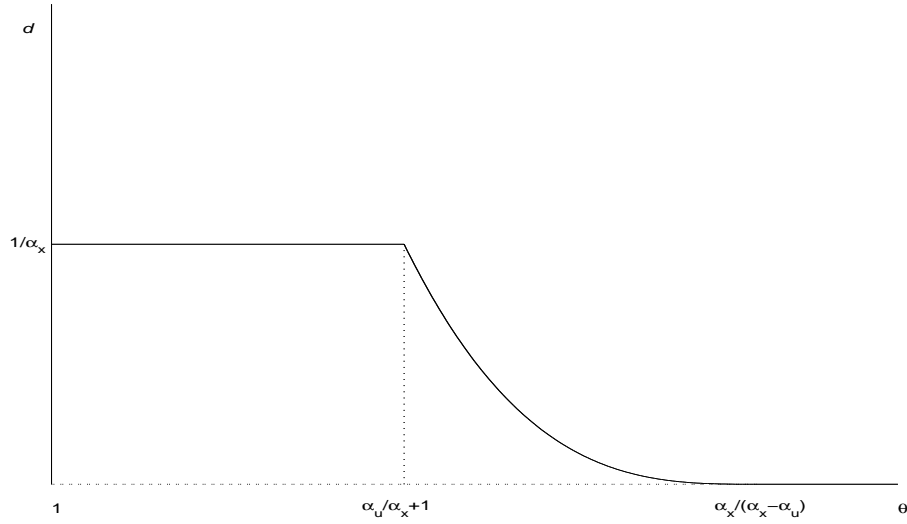


Figure 4: Scenario 3

Note that in the range $\theta \geq \alpha_x/\alpha_u + 1$, the estimator (6) is not consistent and, in particular, d is non-positive. We have chosen to plot $d = 0$ in this range.

Scenario 4: $1 < \alpha_x \leq 2$ and $\alpha_u \geq 2$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{2+\alpha_x}{\alpha_x} \\ \frac{2\theta-(\theta-1)\alpha_x}{2(\theta-1)\alpha_x} & \text{if } \frac{2+\alpha_x}{\alpha_x} < \theta \leq \frac{\alpha_x}{\alpha_x-1} \\ \frac{1}{2} & \text{if } \theta > \frac{\alpha_x}{\alpha_x-1} \end{cases} .$$

See the plot below.

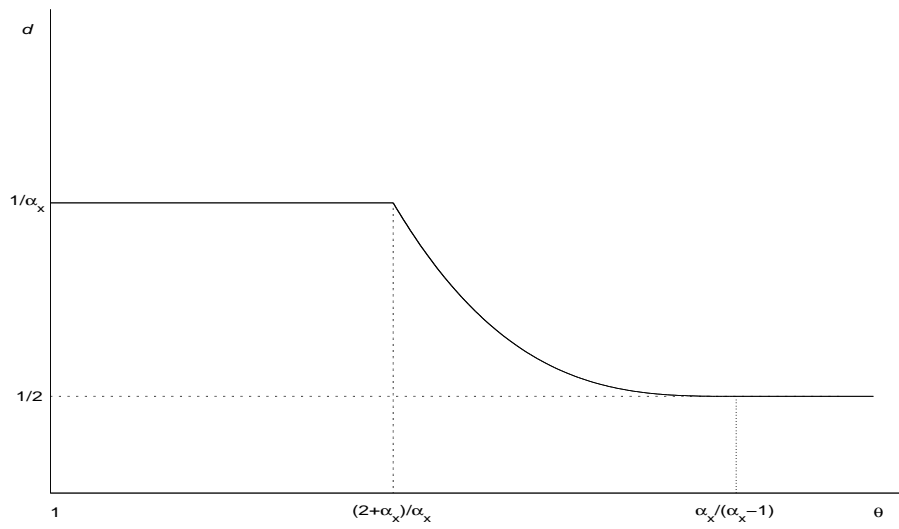


Figure 5: Scenario 4

Scenario 5: $\alpha_x > 1$, $1 < \alpha_u < 2$ and $\alpha_u \leq \frac{\alpha_x}{\alpha_x - 1}$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{\alpha_u}{\alpha_x} + 1 \\ \frac{\theta\alpha_u - (\theta-1)\alpha_x}{(\theta-1)\alpha_u\alpha_x} & \text{if } \frac{\alpha_u}{\alpha_x} + 1 < \theta \leq \frac{\alpha_x}{\alpha_x - 1} \\ 1 - \frac{1}{\alpha_u} & \text{if } \theta > \frac{\alpha_x}{\alpha_x - 1} \end{cases} .$$

See the plot below.

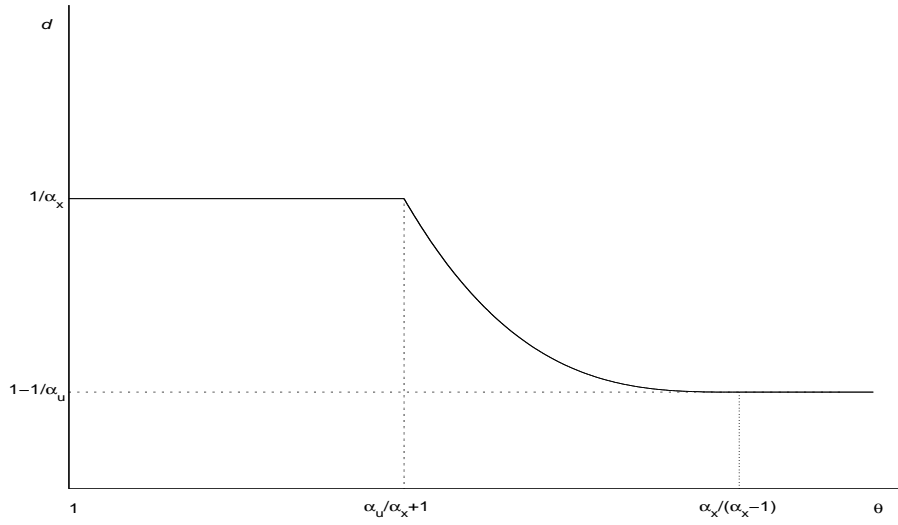


Figure 6: Scenario 5

Scenario 6: $\alpha_x > 2$ and $\alpha_u \geq 2$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{\alpha_x}{\alpha_x - 1} \\ \frac{(\theta - 1)\alpha_x - 1}{(\theta - 1)\alpha_x} & \text{if } \frac{\alpha_x}{\alpha_x - 1} < \theta \leq \frac{2 + \alpha_x}{\alpha_x} \\ \frac{1}{2} & \text{if } \theta > \frac{2 + \alpha_x}{\alpha_x} \end{cases} .$$

See the plot below.

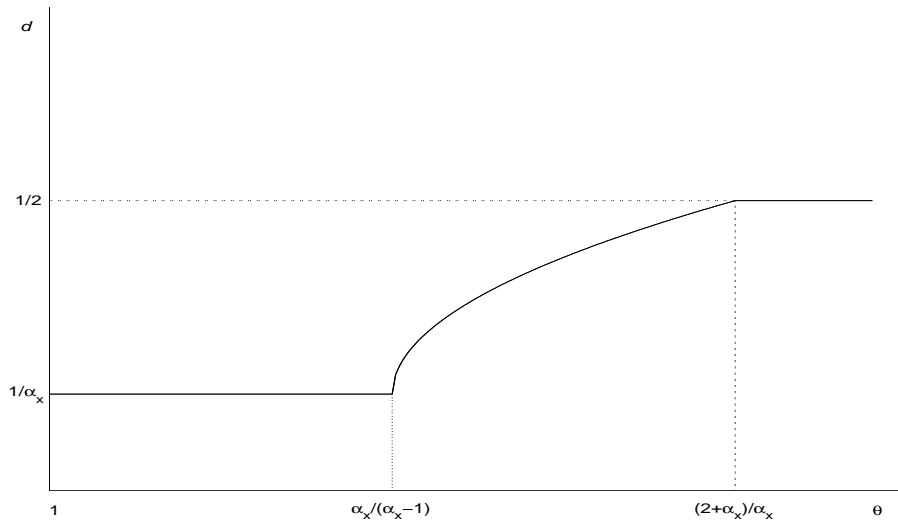


Figure 7: Scenario 6

Scenario 7: $\alpha_x > 2$ and $\frac{\alpha_x}{\alpha_x-1} < \alpha_u < 2$

Here

$$d(\theta) = \begin{cases} \frac{1}{\alpha_x} & \text{if } 1 < \theta \leq \frac{\alpha_x}{\alpha_x-1} \\ \frac{(\theta-1)\alpha_x-1}{(\theta-1)\alpha_x} & \text{if } \frac{\alpha_x}{\alpha_x-1} < \theta \leq \frac{\alpha_u}{\alpha_x} + 1 \\ 1 - \frac{1}{\alpha_u} & \text{if } \theta > \frac{\alpha_u}{\alpha_x} + 1 \end{cases} .$$

See the plot below.

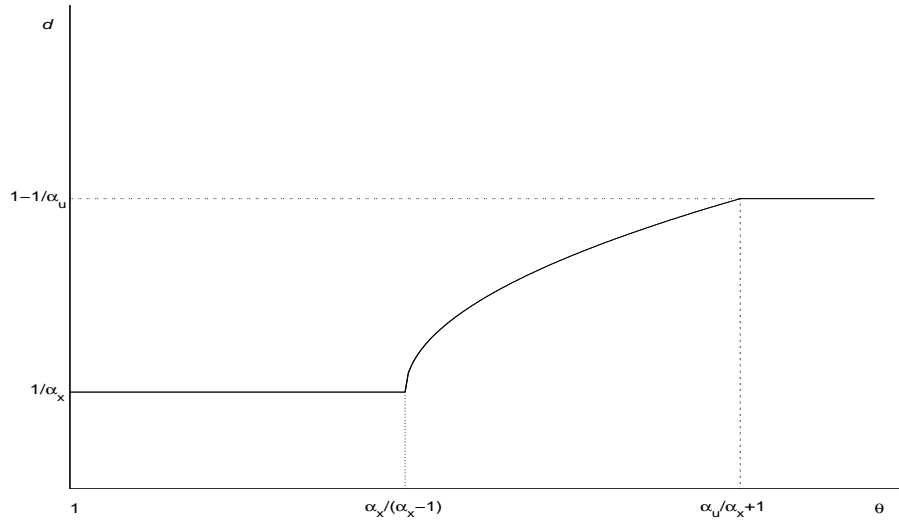


Figure 8: Scenario 7

4 What θ should one use?

The knowledge of α_x and α_u allows us, from the results obtained above, to select the θ that leads to the highest possible rate of decay of Δ_n , ie the highest possible d . In particular, from the plots in the previous section we know that $d(\theta)$ is a *non-increasing* function of θ in Scenarios 1 through 5, and a *non-decreasing* function of θ in Scenarios 6 and 7.

What θ do we choose if the α_x and α_u are unknown or rather we do not know them precisely? This is a common situation since the precision of even the best non-parametric estimators of the tail exponents is not very high; see eg Embrechts et al (1997).

Clearly, the tighter bounds on α_x and α_u we have, the easier it is to select a good θ . In this section we will consider several possible situations. The reader is invited to consider additional possibilities. We will only consider the cases $\alpha_x \geq 1$ and $\alpha_u \geq 1$ here, as those are of relevance in empirical analysis.

Suppose first that we know that

$$\alpha_x \geq 1 \quad \text{and} \quad \alpha_u \geq 2. \quad (138)$$

Then the choice of

$$\theta = 2 \quad (139)$$

always leads to the highest possible rate of decay of Δ_n , ie the highest possible d . Indeed, if $\alpha_x \leq 2$, then Scenario 4 is in force (the boundary case $\alpha_x = 1$ does not distinguish between Scenarios 1 and 4), and, since

$$\frac{2 + \alpha_x}{\alpha_x} = 1 + \frac{2}{\alpha_x} \geq 2,$$

we obtain the optimal $d = 1/\alpha_x$. On the other hand, if $\alpha_x > 2$, then Scenario 6 is in force, and since

$$\frac{2 + \alpha_x}{\alpha_x} = 1 + \frac{2}{\alpha_x} < 2,$$

we obtain the optimal $d = 1/2$.

On the other hand, suppose we know that

$$1 \leq \alpha_x < 2 \quad \text{and} \quad 1 \leq \alpha_u < 2. \quad (140)$$

Then any choice of

$$1 < \theta \leq \frac{3}{2} \quad (141)$$

always leads to the highest possible rate of decay of Δ_n (highest possible d). Indeed, Scenario 5 is in force and

$$\frac{\alpha_u}{\alpha_x} + 1 > \frac{1}{2} + 1 = \frac{3}{2} \geq \theta,$$

and we obtain the highest possible value of $d = 1/\alpha_x$.

Note that in the above cases, and with the choice of θ we are recommending, we will always have $d \geq 1/2$.

Unfortunately, in the range $1 \leq \alpha_u < 2$, if α_x can be bigger than 2, no such efficiency is possible.

To measure the relative efficiency of a given choice of θ , let us introduce the notation

$$R(\theta; \alpha_x, \alpha_u) = \frac{d(\theta; \alpha_x, \alpha_u)}{d^*(\alpha_x, \alpha_u)}, \quad (142)$$

where $d(\theta; \alpha_x, \alpha_u)$ is the value of d corresponding to $\theta, \alpha_x, \alpha_u$ and

$$d^*(\alpha_x, \alpha_u) = \max_{\theta > 1} d(\theta; \alpha_x, \alpha_u). \quad (143)$$

For a set A of (α_x, α_u) let

$$\mathfrak{R}_A(\theta) = \inf_{(\alpha_x, \alpha_u) \in A} R(\theta; \alpha_x, \alpha_u) \quad (144)$$

be the worst efficiency of a given choice of θ . We may then look for a *maxmin* value θ_A such that

$$\mathfrak{R}_A(\theta_A) = \max_{\theta > 1} \inf_{(\alpha_x, \alpha_u) \in A} R(\theta; \alpha_x, \alpha_u) := R_A. \quad (145)$$

If $A \supseteq (2, \infty) \times (1, 2)$, then $R_A = 0$.

Indeed, for a given $\theta > 1$, choose α_x so large that $\alpha_x/(\alpha_x - 1) < \theta$ and let $\alpha_u \downarrow 1$. Then we will be eventually within Scenario 5, and

$$R(\theta; \alpha_x, \alpha_u) = \frac{1 - 1/\alpha_u}{1/\alpha_x} \longrightarrow 0.$$

Hence $\mathfrak{R}_A(\theta) = 0$ for all $\theta > 1$, and so $R_A = 0$, as claimed.

If, however, α_x cannot be arbitrarily large, then things are better.

Let $\alpha^* > 2$, and

$$A = [1, \alpha^*] \times [1, 2). \quad (146)$$

Then any θ in the range

$$1 < \theta \leq \frac{\alpha^{*2} + 2}{\alpha^{*2} - \alpha^* + 2} \quad (147)$$

is a θ_A . Furthermore,

$$R_A = \frac{2}{\alpha^*}. \quad (148)$$

To prove this, consider first the range

$$\theta > \frac{\alpha^*}{\alpha^* - 1}.$$

Here, as in the case $A = (2, \infty) \times (1, 2)$, we see that $\mathfrak{R}_A(\theta) = 0$. Next, we consider the range

$$\frac{\alpha^* + 1}{\alpha^*} < \theta \leq \frac{\alpha^*}{\alpha^* - 1} \quad (149)$$

Note that

$$\inf_{\substack{(\alpha_x, \alpha_u) \times A \\ \alpha_u \leq \alpha_x / (\alpha_x - 1)}} R(\theta; \alpha_x, \alpha_u) = \frac{\theta}{\theta - 1} - \alpha^* \quad (150)$$

and is achieved when $\alpha_x \uparrow \alpha^*$, $\alpha_u \downarrow 1$. On the other hand,

$$\inf_{\substack{(\alpha_x, \alpha_u) \times A \\ \alpha_u > \alpha_x / (\alpha_x - 1)}} R(\theta; \alpha_x, \alpha_u) = \frac{2}{\alpha^*} \quad (151)$$

and is achieved when $\alpha_x \uparrow \alpha^*$, $\alpha_u \uparrow 2$. Therefore, in the range (149)

$$\mathfrak{R}_A(\theta) = \min \left(\frac{\theta}{\theta - 1} - \alpha^*, \frac{2}{\alpha^*} \right) = \frac{2}{\alpha^*} \quad (152)$$

if $(\alpha^* + 1)/\alpha^* < \theta \leq (\alpha_x^2 + 2)/(\alpha_x^2 - \alpha_x + 2)$. Furthermore, $\mathfrak{R}_A(\theta) < 2/\alpha^*$ if $(\alpha_x^2 + 2)/(\alpha_x^2 - \alpha_x + 2) < \theta \leq \alpha^*/(\alpha^* - 1)$.

Similarly in the range

$$1 < \theta \leq \frac{\alpha^* + 1}{\alpha^*}, \quad (153)$$

we obtain, in a similar way, that

$$\mathfrak{R}_A(\theta) = \frac{2}{\alpha^*}, \quad (154)$$

Therefore, both (147) and (148) follow. In this situation we can guarantee $d \geq 2/(\alpha^*)^2$ with the choice of θ recommended above.

The above discussion of the ways to select the parameter θ focuses on the rate of convergence to the true value, which is, clearly, the single most important criterion. With the rate of convergence kept fixed, however, other things become important. Among them is the spread of the limiting distribution. To compare such spreads and, hence, to be able to tell more about good ways to select θ , we performed a simulations study.

Design of simulation. From the viewpoint of empirical evidence, we consider $\alpha_x \in [1, 2)$ and $\alpha_u \in [1, 2)$. To implement data-generating processes, we have selected $\alpha_x, \alpha_u = 1, 1.2, 1.4, 1.6, 1.8, 1.99$.⁴ For sample size we choose $n = 50, 100, 250, 500, 1,000, 2,000, 5,000, 10,000$ and ∞ , where the limiting distributions are calculated from Scenario 5. We use a length of quantile $\xi_{0.975} - \xi_{0.025}$ as a spread measure, where ξ_p is the p th quantile of the simulated distribution of (7).⁵ According to the recommendation in (141), we use $\theta \in (1, 1.5]$. For implementation we have selected $\theta = 1.05, 1.1, 1.15, 1.2, 1.25, 1.3, 1.35, 1.4, 1.45, 1.5$. To determine simulated densities for each estimate, 10,000 replications were made. Figure 9 shows the θ minimizing the spread for selected α_x, α_u and sample sizes.

The selected θ shows noticeable irregularity, even for large samples. Nevertheless, some useful rules for choosing the θ can be formulated as

$$\theta = \begin{cases} \alpha_u & \text{if } \alpha_u < 1.5 \\ 1.5 & \text{if } \alpha_u \geq 1.5 \end{cases}$$

Here, the parameter α_u plays a key role, while the role of α_x and the sample size seem to be less important.

⁴In order to better see the behavior of the estimates near the boundary points 1 and 2, a more detailed selection was used in another simulation. The results show that the transition from 1 to numbers bigger than 1 (1.01 and 1.05 were additionally chosen in the simulation) and from 2 to numbers smaller than 2 (1.99 and 1.95) is smooth.

⁵The results are very robust against taking other quantiles and other spread measure such as variation; see Samorodnitsky and Taqqu (1994, Ch 2) for an explanation of the variation.

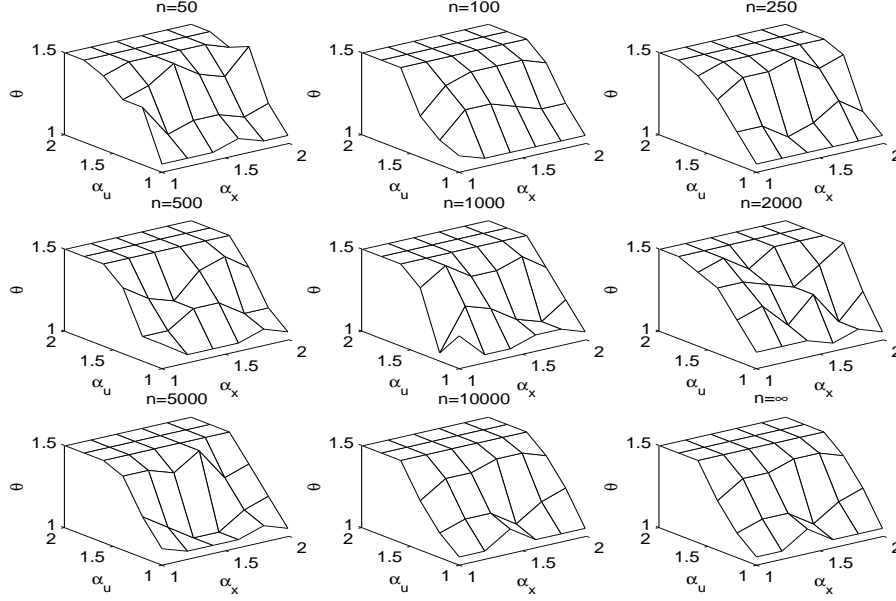


Figure 9: θ for selected α_x and α_u

5 Limiting and finite-sample distributions of Δ_n under Scenario 5

In this section we consider the most interesting scenario in practice, Scenario 5. Since the limiting distribution in (51) is a non-standard one, we perform simulations and response surface analysis and give both the limiting distributions by given α_x and α_u , and finite-sample distributions.

5.1 Limiting distributions of Δ_n

The behavior of the limiting distribution in (51) based on Scenario 5 is numerically analyzed. Figure 10 shows simulated limiting distributions for selected α_x and α_u . For this simulation we used θ according to the choice rule described above.

A comparison of the limiting distributions for various α_x and α_u shows that they become more dispersed if α_u decreases or α_x increases. For example, 0.0113, 0.0032, 0.0017 for $\alpha_u = 1.0, 1.5, 1.99$, respectively, for a given $\alpha_x = 1.5$, and 0.0366, 0.0032, 0.0001 for $\alpha_x = 1.99, 1.5, 1.0$, respectively, for a given $\alpha_u = 1.5$. For the spread of the limiting distributions, α_x plays a more important role than α_u .

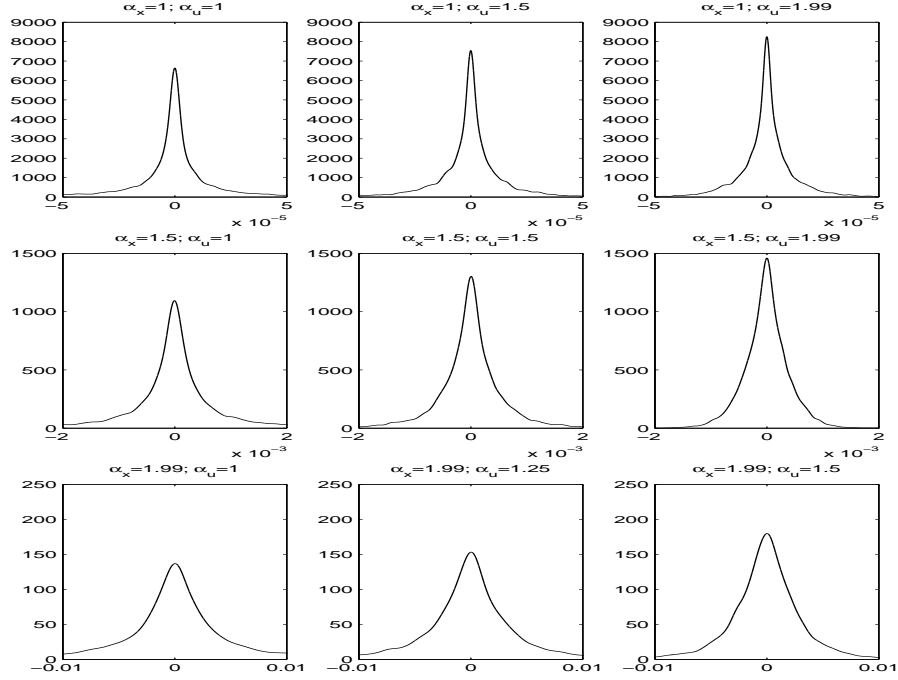


Figure 10: Simulated limiting distributions for selected α_x and α_u

Rather than simply tabulating critical values for a few selected sample sizes, α_x -values and α_u -values, we employ response surface techniques to present our simulation results in a compact fashion. In addition, this approach allows us to derive approximate critical values for wide ranges of α_x - and α_u -values and facilitates computational implementation. Response surface methodology has been used in various statistical and econometric applications (see Myers et al 1989).

Response surface analysis was applied to approximate selected quantiles of the limiting distributions Δ_n given in (51) generated from the 10,000 replications. Specifically, we focused on the 50, 60, 70, 80, 90, 95 and 99% quantile values and, hence, for the fitting of a response surface, on the 175 (α_x, α_u, q) -combinations considered. For the surface a functional form was specified, a high-order polynomial in α_x , α_u and q . Specifically, we estimated

$$\xi(\alpha_x, \alpha_u, q) = \sum_{h=1}^4 \sum_{i=2}^3 \sum_{j=0}^2 a_{h,i,j} \alpha_x^{4i} \alpha_u^{-2j} (1-q)^{-h/2} + \varepsilon_{\alpha_x, \alpha_u, q},$$

$$\alpha_x, \alpha_u \in [1, 2); \quad q \in [0.5, 0.99].$$

To derive the approximate response surface, we selected the subset of regressors which

maximized the adjusted- R^2 value. To take possible heterogeneity of the approximation error $\varepsilon_{\infty, \alpha_x, \alpha_u, q}$ into account, the generalized least squares method was used for estimation, although there did not seem to be systematic heterogeneity due to variations of the characteristic exponent α . The estimation results are reported in Table 1.

Table 1.
Estimated coefficients (a_{hij}) in the response surface function^a

j	h i	1		2		3		4	
		2	3	2	3	2	3	2	3
1		-443.1 (-0.97)		309.7 (1.34)	-55.32 (-1.17)	-33.61 (-1.06)		2.6767 (2.14)	-1.6330 (-4.79)
2					38.51 (2.42)		-8.0838 (-2.85)	0.1374 (1.65)	0.5719 (3.90)
3								-0.0019 (-3.21)	0.0020 (3.12)

^aThe estimates are multiplied by 10^6 . t -values are given in brackets.

Various measures of fit, namely $R^2 = 0.9991$, adjusted- $R^2 = 0.9990$,⁶ $\hat{\sigma}_\varepsilon = 0.0050$, mean $|\hat{\varepsilon}| = 1.7706 - 10^{-3}$, and $\max |\hat{\varepsilon}| = 0.0398$ indicate adequate fits. As expected, the absolute goodness of fit deteriorates as the significance level decreases, especially when α_x and α_u approach 2.⁷

⁶The negligibility of the constant terms justifies the use of the adjusted- R^2 value for selecting the regressor subset. Its use would be inappropriate if the zero restrictions on the constant terms did not hold.

⁷To obtain better response surface approximations, simulations with additional intermediate values for α_x and α_u could have been conducted. However, the surfaces turn out to be rather smooth with respect to both variables, which means additional simulations would have led to only negligible improvements in the response surface approximations. Selected comparisons between simulations with intermediate (α_x, α_u) -combinations and corresponding response surface approximations support this conclusion.

5.2 Finite-sample distributions of Δ_n

To examine the finite-sample behavior of Δ_n , we simulated samples of length $n = 50, 100, 250, 500, 1,000, 2,000, 5000$ and $10,000$. Both regressors x_t and disturbances u_t were drawn from symmetric standard stable Paretian distributions (ie $c = 1$ and $\delta = 0$) with the characteristic exponent α_x and α_u assuming values $1.0, 1.25, 1.5, 1.75, 1.99$, θ having been chosen by the rule in the previous subsection. For each of the 200 possible (α_x, α_u, n) -combinations, 10,000 replications with the same seed were generated with the algorithm of Weron (1996) as a modified version of Chambers et al (1976).⁸

Overall properties of the distributions for finite samples with respect to α_x and α_u are similar to those of the limiting distributions. For supplying critical values for finite samples, response surface analysis was again applied to approximate selected quantiles of Δ_n generated from the 10,000 replications. Specifically, we focused on the 1, 5 and 10% critical values and, hence, for the fitting of a response surface on the 600 $(n, \alpha_x, \alpha_u, q)$ -combinations considered. For the surface a functional form was specified, a high-order polynomial in n, α_x and α_u accompanied by quantiles. Specifically, we estimated the following regression and the estimated coefficients of the response surface regression are summarized in Table 2.

$$\xi(n, \alpha_x, \alpha_u, q) = \sum_{h=1}^5 \sum_{i=1}^4 \sum_{j=2}^3 a_{h,i,j} n^{-h/2} \alpha_x^{4i} \alpha_u^{-2j} (1-q)^{-1} + \varepsilon_{n,\alpha_x,\alpha_u,q},$$

$$n \in [50, 10000]; \alpha_x, \alpha_u \in [1, 2]; q \in [0.9, 0.99].$$

Various measures of fit, namely $R^2 = 0.9964$, adjusted- $R^2 = 0.9962$, $\hat{\sigma}_\varepsilon = 0.1034$, mean $|\hat{\varepsilon}| = 0.0613$ and max $|\hat{\varepsilon}| = 0.7816$ indicate adequate fits. As expected, the absolute goodness of fit deteriorates as the significance level decreases, especially when α_x and α_u approach 2.

Next, we use a Kolmogorov-Smirnov test to check how close the finite-sample distributions of Δ_n are to the limiting distributions. Figure 11 shows the QQ plots of the empirically estimated distributions F_n against the limiting distribution F_0 (51) with

⁸This algorithm is analytically identical with the widely used stable random number generator written by J.H. McCulloch which is available on McCulloch's website: <http://economics.sbs.ohio-state.edu/jhm/jhm.html>

Table 2.

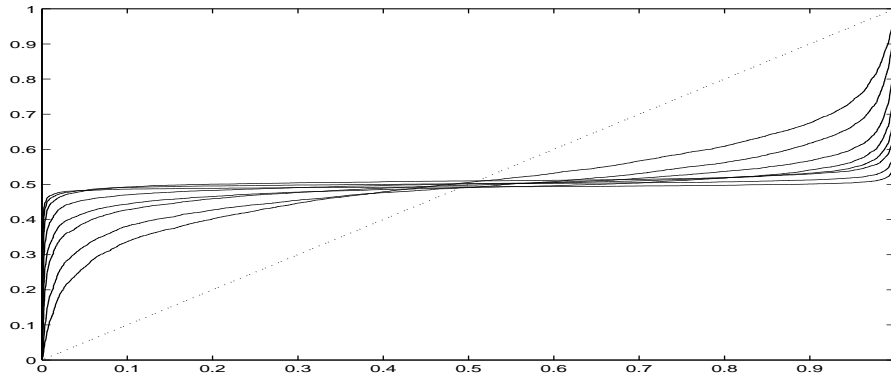
Estimated coefficients (a_{hij}) in the response surface function

h	i	1		2		3		4	
		2	3	2	3	2	3	2	3
1					0.0103 (2.6127)			3.5-10 ⁻⁵ (2.1876)	
2			0.6069 (1.1815)			0.1480 (1.5720)	-0.2907 (-2.6026)	-0.0130 (-2.0952)	0.0221 (3.2541)
3		9.6392 (2.4731)	-8.5434 (-1.3986)	7.2154 (1.0088)		-5.4666 (-1.4808)	9.0033 (2.3214)	0.4135 (1.7690)	-0.7028 (-2.8803)
4				-127.5576 (-1.1743)	13.242 (1.0160)	60.999 (1.3993)	-94.389 (-2.1930)	-4.3110 (-1.6067)	7.3192 (2.6787)
5				433.11 (1.0254)		-199.83 (-1.2848)	307.22 (2.0620)	13.970 (1.4796)	-23.981 (-2.5249)

$j = 50,000$, for $\alpha_x = 1.5$ and $\alpha_u = 1.5$, where $n = 50, 100, 250, 500, 1,000, 2,000, 5,000$ and $10,000$. The dotted 45° line from (0,0) to (1,1) corresponds to the limiting distribution. The empirical values of the Kolmogorov-Smirnov (KS) test statistic:

$$\sup_x |F_n(x) - F_0(x)| \quad (155)$$

are obtained by taking the largest deviations from the dotted line.

Figure 11: Simulated limiting distributions for selected $\alpha_x = 1.5$ and $\alpha_u = 1.5$

Note that the critical value at the 95% significance level is $1.3851/\sqrt{n}$. As is clearly shown in Figure 11, the finite-sample distributions differ significantly from

the limiting distributions even at $n = 10,000$. For all cases the null hypothesis that $F_n = F_0$ cannot be accepted. For various α_x and α_u the KS test shows the same result. This means that for empirical work the limiting distributions is only a poor approximation of the finite sample distributions, which means we need tables of critical (quantile) values of each combination of (α_x, α_u) for a range of sample sizes n .

6 Concluding remarks

One can see that blindly using the OLS approach $\theta = 2$ can lead to very inefficient estimators of the regression coefficient. A much better approach is to take the tails into account. Even if the tails of the regressors and disturbances are known only approximately, this can still provide valuable information for selecting a good value of θ and, hence, constructing a more efficient estimator. Iterated procedures in which the tails and the regression coefficient are estimated simultaneously should be considered.

JEONG-RYEOL KURZ-KIM

email: jeong-ryeol.kurz-kim@bundesbank.de

Deutsche Bundesbank

Wilhelm-Epstein-Strasse 14, D-60431 Frankfurt am Main, Germany

SVETLOZAR T. RACHEV

email: rachev@pstat.ucsb.edu

Institut für Statistik und Mathematische Wirtschaftstheorie

Universität Karlsruhe, Kollegium am Schloss, D-76128 Karlsruhe, Germany

and

Department of Statistics and Applied Probability

University of California, Santa Barbara, CA 93106, USA

GENNADY SAMORODNITSKY

email: gennady@orie.cornell.edu

School of Operations Research and Industrial Engineering, Cornell University

Ithaca, NY 14853, USA

References

- Bachelier, L (1900): Théorie de la spéculation, *Annales Scientifiques de l'Ecole Normales Supérieure* III-17:21–86. Translated in *The Random character of stock market prices*, Ed. P.H. Cootner, 17–78, Cambridge, MIT Press, 1964.
- Billingsley, P. (1999): *Convergence of Probability Measures*, Wiley, New York, 2nd edition.
- Blattberg, R. and Sargent, T. (1971): Regression with Non-Gaussian Stable Disturbances: Some Sampling Results, *Econometrica* 39, 501–510.
- Chambers, J.M., C.L. Mallows and B.W. Stuck (1976): A method for simulating stable random variables, *Journal of the American Statistical Association* 71, 340–344.
- Embrechts, P., C. Klüppelberg and T. Mikosch (1997): *Modelling Extremal Events for Insurance and finance*, Springer-Verlag, Berlin.
- Fama, E.F. (1965): The Behavior of Stock Market Prices, *Journal of Business* 38, 34–105.
- Feller, W. (1966): *An Introduction to Probability Theory and Its Applications*, Wiley, New York.
- Kim, J.-R. (2000): *Eine Verallgemeinerte Kapitalmarkttheorie*, Unpublished Habilitationsschrift.
- Kwapień S. and N. Woyczyński (1992): *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, Boston.
- Laha, R. and V. Rohatgi (1979): *Probability Theory*, Wiley, New York.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices, *Journal of Business* 36, 394–419.
- Myers, R.H., I. Khuri and W.-H. Carter, Jr. (1989): Response surface methodology: 1966-1988, *Technometrics* 31, 137–157.

- Rachev, S.T., J.-R. Kim and S. Mittnik (1999): Stable Paretian econometrics. Part I and II, *The Mathematical Scientist* 24, 24–55 and 113–127.
- Rachev, S.T. and S. Mittnik (2000): *Stable Paretian Models in Finance*, Wiley.
- Resnick, S.I. (1987): *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, New York.
- Samorodnitsky, G. and M.S. Taqqu (1994): *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York.
- Weron, R. (1996): On the Chambers-Mallows-Stuck method for simulating skewed stable random variables, *Statistics and Probability Letters* 28, 165–171.
- Zolotarev, V.M. (1986): *One-dimensional Stable Distributions*, Translation of Mathematical Monographs, American Mathematical Society, Vol. 65, Providence.